DOCUMENT RESUME

ED 422 331		TM 028 877
AUTHOR	Barton, Paul E.; Coley, Richard J.	
TITLE	Growth in School: Achievement Gains from the Fourth to the Eighth Grade. Policy Information Report.	
INSTITUTION	Educational Testing Service, Princeton, NJ. Policy Information Center.	
PUB DATE	1998-05-00	
NOTE	21p.	
AVAILABLE FROM	Policy Information Center, Mailstop 04-R, Educational Testing Service, Rosedale Road, Princeton, NJ 08541-0001; e-mail: pic@ets.org; World Wide Web: http://www.ets.org/research	
PUB TYPE	Reports - Descriptive (141)	
EDRS PRICE	MF01/PC01 Plus Postage.	
DESCRIPTORS	*Achievement Gains; *Cohort Analysis; *Educational Trends; Elementary Education; *Elementary School Students; Grade 4; Grade 8; International Studies; *Mathematics Achievement; National Surveys; Science Education	
IDENTIFIERS	*National Assessment of Educational Progress; Third International Mathematics and Science Study	

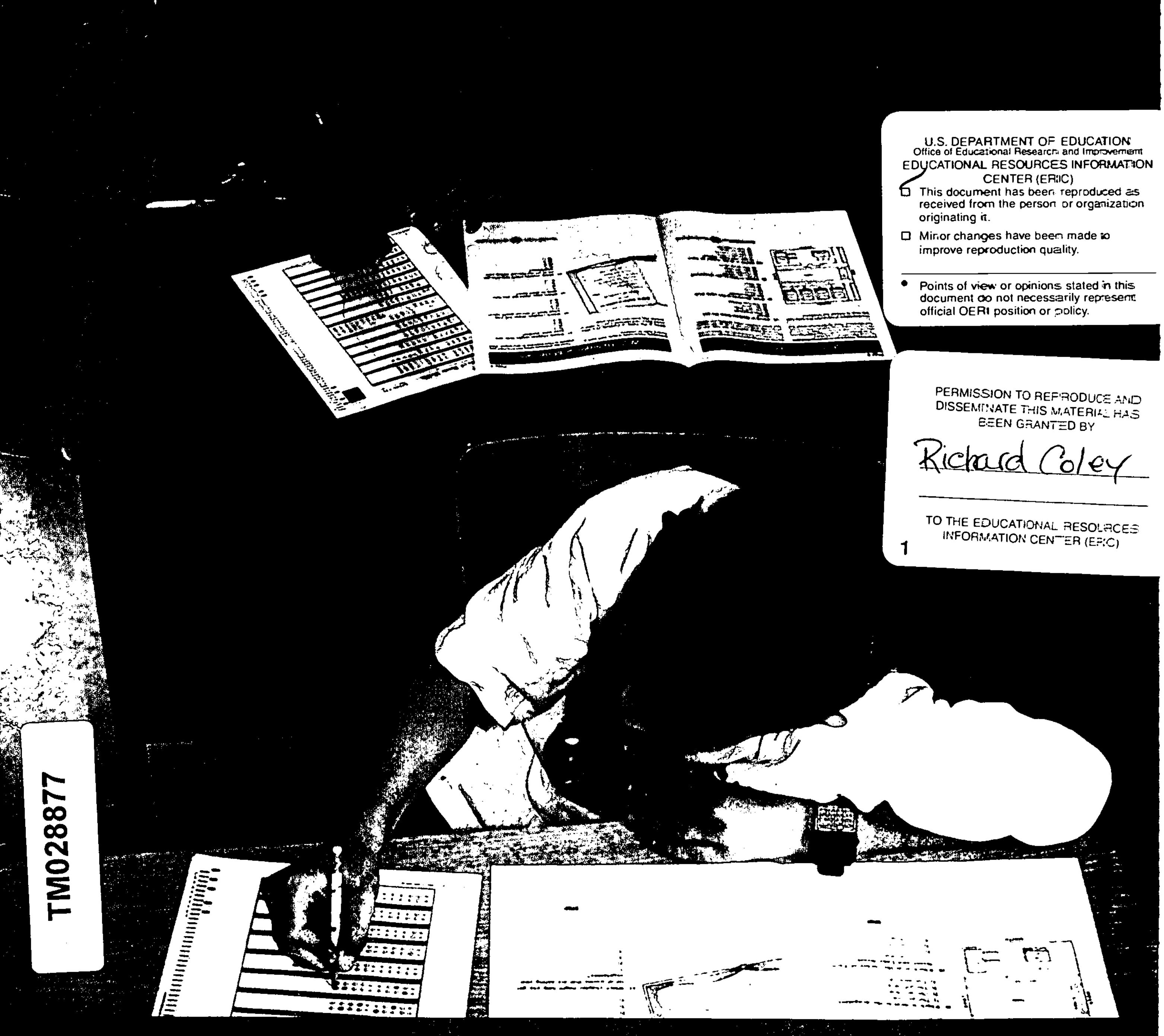
ABSTRACT

The Third International Mathematics and Science Study disclosed that while U.S. students did well in the fourth grade in comparison with students from other countries, they had slipped considerably by eighth grade. This study was conducted to see what could be learned about achievement growth between grades 4 and 8. Achievement growth was investigated using the National Assessment of Educational Progress (NAEP), which has been redesigned so that it is possible to track cohorts of students and to determine the value-added in terms of education between fourth and eighth grades. When NAEP cohort records were examined, it was found that the average NAEP scores of students are slightly higher today than those of students of 20 or 25 years ago, but the same is not true of cohort growth between grades 4 and 8. Cohort growth is the same as, or lower than, it was for the earliest period for which data are available. When individual states are studied, there is little cohort growth between the fourth and eighth grades. Measuring and examining cohort growth has the potential to provide a new and important dimension in understanding trends in educational achievement. Research must then determine the factors related to cohort growth changes. (Contains 5 tables and 11 figures.) (SLD)

POLICY INFORMATION REPORT

Growth in School:

Achievement Gains from the Fourth to the Eighth Grade



Educational Testing Service

POLICY INFORMATION CENTER
Research Division
Educational Testing Service, Princeton, New Jersey 08541-0001

CONTENTS

Preface 2

Acknowledgments 2

Framing the Question 3

Measuring Cohort Growth 5

Trends in Cohort Growth 6

Some Cohort Growth Comparisons 8

State Comparisons 10

Describing and Understanding Cohort Growth 12

In Conclusion 15

This report was written by:

Paul E. Barton and
Richard J. Coley
Policy Information Center
Educational Testing Service

The views expressed in this report are those of the authors and do not necessarily reflect the views of the officers and trustees of Educational Testing Service.

Additional copies of this report can be ordered for $9.50 (prepaid) from:

Policy Information Center
Mailstop 04-R
Educational Testing Service
Rosedale Road
Princeton, NJ 08541-0001
(609) 734-5694
Internet - pic@ets.org
The report can also be downloaded from
http://www.ets.org/research

May 1998

PREFACE

We wrote this report for two reasons. First, the Third International Mathematics and Science Study (TIMSS) disclosed that while U.S. students did well in the fourth grade in comparison with students from other countries, they had slipped considerably by the eighth grade. We wanted to examine data from the National Assessment of Educational Progress (NAEP) to see what could be learned about achievement growth between the fourth and eighth grades in the U.S.

The second, and related, reason is that the redesign of NAEP in 1984 made it possible to track cohorts of students from age 9 to 13, or from grades 4 to 8. Enough time has now passed under this new design that we can now use it to compare groups of students and individual states in terms of "value added" between the fourth and eighth grade. In addition, we can see whether the "value added" has increased, stayed the same, or decreased over time.

We believe that this approach is important; perhaps as important as the normal NAEP approach of seeing how achievement changes for students in the same grade or of the same age, over a period of time. We see an advantage is looking at NAEP data *both* ways.

While we have dealt exclusively with data on students in the fourth and eighth grades (or ages 9 and 13), we note that TIMSS is reporting twelfth-grade results as this report goes to press. The news was that U.S. twelfth-graders did not perform well in the international comparison; in fact, they performed less well than U.S. eighth-graders had in a previous TIMSS assessment.

Extending our analysis of NAEP data to the twelfth grade, however, poses special problems. One is that some students drop out of school between grades 8 and 12, so the student samples can be different. Another problem is that in trend comparisons, dropout rates may be higher or lower during different time periods, making comparability problematic.

For these reasons, along with the data made available by the redesign of NAEP, we focus this report on the "value added" in achievement between the fourth and eighth grades.

Paul E. Barton
Richard J. Coley

ACKNOWLEDGMENTS

The data in this report are from the National Assessment of Educational Progress, a congressionally mandated project of the National Center for Education Statistics, U.S. Department of Education.

David Freund, of Educational Testing Service provided statistical analysis of the NAEP data. John Mazzeo and Howard Wainer of Educational Testing Service provided reviews of this report, as did Cynthia Prince of the National Education Goals Panel. These reviewers do not necessarily agree with all the views and interpretations in this report.

Carla Cooper provided cover design and desktop publishing services. Kirsty Brown was the editor. Jim Chewning coordinated production.

FRAMING THE QUESTION

There has been an explosion in standardized testing over the last 20 years or so, both to measure individual student progress and to monitor achievement at the school, district, state, or national levels. Students are almost always grouped by grade level, and the traditional focus is on tracking how students in those grade levels compare over time. Policy makers ask such questions as: How do today's fourth-graders compare to fourth-graders 10 years ago in mathematics? In this report, such statistics are called "average score trends."

One of the major tools for accomplishing this monitoring is the National Assessment of Educational Progress, or NAEP, as it is usually called. This "nation's report card" is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. Begun in 1969, NAEP reports achievement at grades 4, 8, and 12 (or for ages 9, 13, and 17 for long-term trends). NAEP reports focus, for example, on how today's 9-year-olds compare to 9-year-olds at some prior time.

During the 1990s, NAEP has conducted voluntary state assessments. Most states participate, although not always the same states in any one assessment year. In sum, NAEP is recognized as a reliable and well-understood means of viewing the performance of students and education systems.

NAEP is traditionally used to describe average score trends in various subjects. The most recent NAEP trend report summarizes the findings over the past 20 years or so:

> *In general, the trends in science and mathematics show early declines or relative stability followed by improved performance. In reading and writing, the results are somewhat mixed; although some modest improvement was evident in the trends for reading assessments, few indications of positive trends were evident in the writing results.*[1]

NAEP results can also be viewed in ways that can help answer different questions. Recently, national attention has focused on the results of the Third International Mathematics and Science Study (TIMSS), conducted in 1995, comparing fourth- and eighth-graders in mathematics and science. TIMSS tested half a million eighth-grade students in 41 countries, in 30 different languages. At the fourth-grade level, 26 countries took part.

At the fourth-grade level, the news was good for the U.S. in both math and science. Only one country, Korea, outperformed U.S. students in science, and U.S. scores were above the international average in mathematics. However, the results were not as favorable at the eighth-grade level. While the U.S. scored above the international average in science, it outperformed just 15 of the 40 countries. And in mathematics, the U.S. scored *below* the international average, outperforming just seven of the 40 countries.

[1] J.R. Campbell, K. E. Voelkl, and P. L. Donahue, *NAEP Trends in Academic Progress*, Washington, DC: National Center for Education Statistics, 1997.

One question that emerged was why did the U.S. slip so much in the international comparisons between the fourth and eighth grade? The traditional way of viewing NAEP results (examining average score trends) does not answer this question. William H. Schmidt and his colleagues conducted an examination of the curriculum in each country, and how those curricula compare at the fourth and eighth grades. Schmidt concluded that what is taught in U.S. mathematics classes at the eighth grade is less advanced and less focused than the curricula of other countries included in TIMSS.

The related question addressed in this report concerns what light NAEP data can shed on this matter of student progress from the fourth to the eighth grade.

Measuring Cohort Growth

The NAEP data base can be used to examine cohort growth and how that growth has changed over time. In the redesign of NAEP by the Educational Testing Service, beginning with the 1984 assessment, a conscious effort was made to enable NAEP to track the educational achievement of the *same cohort* of students. This was done by spacing the grade or age levels assessed four years apart (fourth, eighth and twelfth grade; or age 9, 13, or 17), and conducting the assessment in each subject at least every four years.

While this assessment pattern has not always been followed precisely, it has been used closely enough to permit some comparisons based on following a *cohort* of students during a four-year period of schooling.

Another key feature of the redesign that allows us to follow cohorts is the use of a *single* scale of achievement (from 0 to 500), that encompasses students at all three grade or age levels. Comparing the progress of cohorts of students was one of the reasons for moving to this developmental scale.

To measure NAEP cohort growth, we look at the average scores of 9-year-olds, and then look at the scores of the same cohort of students four years later when they are 13. For example, the average score in mathematics for 9-year-olds in 1978 was 219; the average score four years later in 1982 when they were 13 was 269. *In the four years, there was cohort growth of 50 points on the mathematics scale.*

In 1992, the average math score for 9-year-olds was 230. In 1996, the average for 13-year-olds was 274. *The cohort growth during the more recent four years was 45 points (after rounding).*

As these data show, there was a loss in the cohort growth between age 9 and 13 over the two time periods. The *loss* was five scale points of growth and was statistically significant (from 50 scale points between 1978 and 1982 to 45 scale points between 1992 and 1996).

Thus, when we look at the change in cohort growth, we get a different picture from when we look at average score trends. The scores of 9-year-olds reflect what happens to the child from birth to kindergarten and then from grades 1 to 4. Over time, children may develop more or less from birth to age 9, quite independently of what happened during the fourth grade, or what happened in school from the first to the fourth grade. Changes in the scores of 9-year-olds may result from what happened before they started school, in the home or in the community.

However, when we measure cohort growth as students move from age 9 to age 13, we are getting much closer to what has happened solely due to their schooling, although even here outside influences can affect the value added in these scores differently in the time periods being compared.

The news is not encouraging — between 1978 - 1982 and 1992 - 1996 students showed less score gain in mathematics from fourth to eighth grade, not more.

7

Trends in Cohort Growth

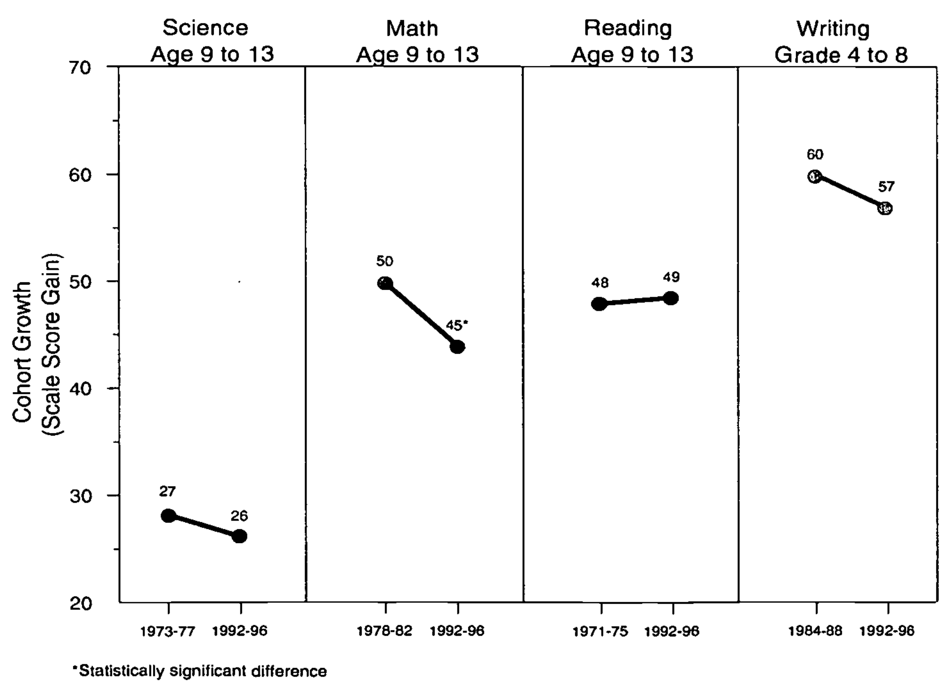
The trends in cohort growth for science, mathematics, reading, and writing are shown in Figure 1. Growth is down in math and is stable in science, reading, and writing (the small changes shown in Figure 1 are not statistically significant, except for the change in math). There were *no* increases in cohort growth over the four subjects from age 9 to age 13.

Again, this is a different view from that normally provided by NAEP analysts. While, in general, there were increases in average score trends over this time period, when 9-year-olds and 13-year-olds are compared to their counterparts in past years in science, math, and reading, there has been no cohort growth. The average score trend gains we have seen are those made by age 9, and carried through to ages 13 and 17— or partially lost as students continue on in school.

Cohort growth from age 9 to 13 can also be examined by race, for White and Black students.[2] The results are shown in Figure 2. Only one of the cohort growth differences shown is statistically significant: the drop in math score gain between age 9 and 13 for White students. Of course, when viewed in terms of average age- or grade-based NAEP scores, achievement levels differ considerably by race and ethnicity.

As the data have shown, the trends in cohort change over four years of schooling can

Figure 1: Trends in NAEP Cohort Growth



Source: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep.

[2] Due to variations in immigration patterns over time, this approach may not be valid for Hispanic students and is not included here.

be quite different from average score trends for fourth- and eighth-grade students. A summary of these comparisons is provided in Table 1 for all 9- and 13-year-olds (or fourth- and eighth-graders), and in Table 2 for Black students. Generally, the cohort growth trends from the fourth to the eighth grade have been level or down, while the average score trends of fourth- and eighth-graders have been almost always up in a comparable period of time.

Table 1: Trends in Cohort Growth Compared to Average Score Trends for 9- and 13-year-olds*

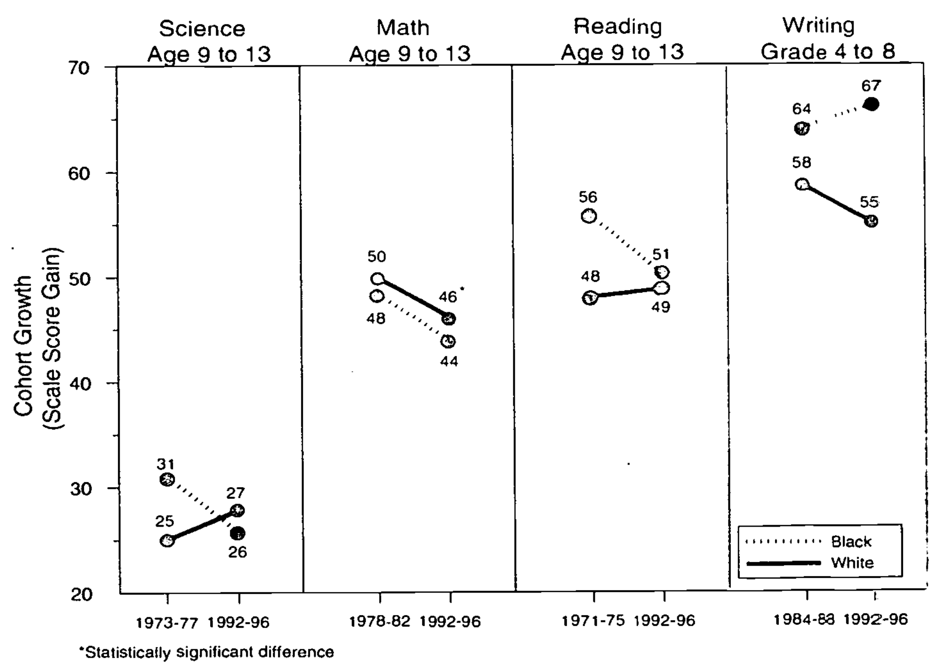
	Cohort Growth, Age 9 to 13	Average Score Trend, Age 9	Average Score Trend, Age 13
Science	Level	Up	Up
Mathematics	Down	Up	Up
Reading	Level	Up	Up
Writing**	Level	Level	Level

Table 2: Trends in Cohort Growth Compared to Average Score Trends for Black 9- and 13-year-olds*

	Cohort Growth, Age 9 to 13	Average Score Trend, Age 9	Average Score Trend, Age 13
Science	Level	Up	Up
Mathematics	Level	Up	Up
Reading	Level	Up	Up
Writing**	Level	Level	Level

Source for Tables 1 and 2: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center.
See http://nces.ed.gov/naep. "False Discovery Rate" procedure used to test for significance.
*Science cohort changes are from 1973-77 to 1992-96. Average science score trends are from 1973 to 1996. Mathematics cohort changes are from 1973-77 to 1992-96. Average mathematics score trends are from 1973 to 1996. Reading cohort changes are from 1971-75 to 1992-96. Average reading score trends are from 1971 to 1996. Writing cohort changes are from 1984-88 to 1992-96. Average writing score trends are from 1984 to 1996.
**Writing was administered to fourth- and eighth-graders.

Figure 2: Trends in NAEP Cohort Growth, by Race



Source: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep.

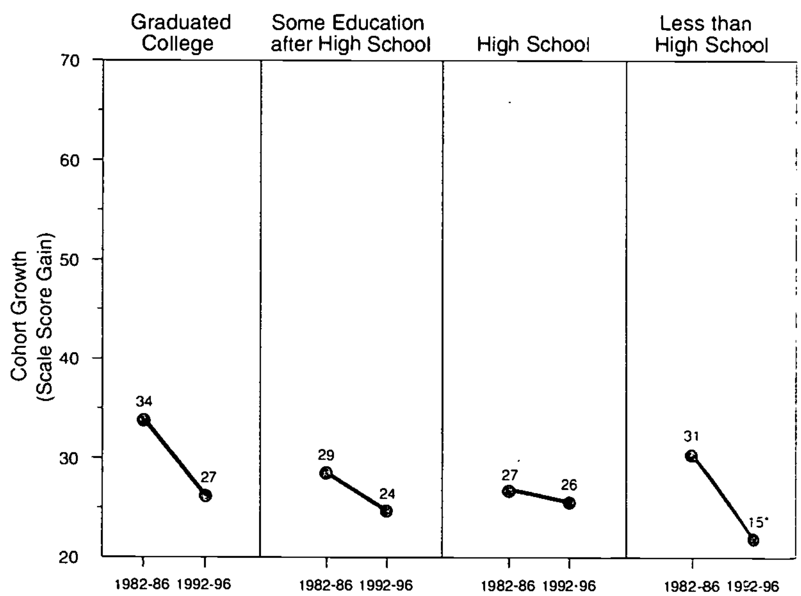
Some Cohort Growth Comparisons

Cohort growth in the NAEP science assessment was examined for students with different levels of parent education, for the periods 1982-86 and 1992-96 (see Figure 3). Cohort growth was stable for students whose parents had a high school education, and for those with more than a high school education. However, the cohort growth was cut in half for students whose parents had less than a high school education, dropping from 31 scale points to 15, a statistically significant difference.

Other comparisons can be made simply by using cohort growth changes in the latest period for which they are available. In Table 3, we can see that the highest cohort growth in mathematics between the fourth and eighth grade, from 1992 to 1996, was among students whose parents had graduated from college (+55 points). That difference was statistically significant, as was the cohort growth for students whose parents had graduated from high school. We are less sure about these comparisons than others in this report, however, since students' reports of parent education can be inaccurate. While this is the same cohort of students in grade 4 and 8, the eighth-graders may have reported parent education more accurately than when they were fourth-graders.

In terms of race/ethnicity, the small differences in cohort growth were not statistically significant (see Table 4). Likewise, during the same period, small differences in cohort growth across regions of the country were not statistically significant (see Table 5).

Figure 3: Trends in NAEP Cohort Growth in Science, by Parent Education



*Statistically significant difference

Source: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep.

Table 3: NAEP Math Scores and Cohort Growth, by Parent Education			
	4TH GRADE, 1992	8TH GRADE, 1996	COHORT GROWTH
All Students	220	272	+52
Students who reported the parents' highest level of education as. . .			
Did Not Finish High School	205	254	+49
Graduated from High School	215	261	+46*
Some Education After High School	225	279	+54
Graduated from College	227	282	+55*

Table 4: NAEP Math Scores and Cohort Growth, by Race/Ethnicity			
	4TH GRADE, 1992	8TH GRADE, 1996	COHORT GROWTH
All Students	220	272	+52
Students who indicated their race/ethnicity as . .			
White	228	282	+54
Black	193	243	+50
Hispanic	202	251	+49
Asian/Pacific Islander	232	---	---
American Indian	211	264	+53

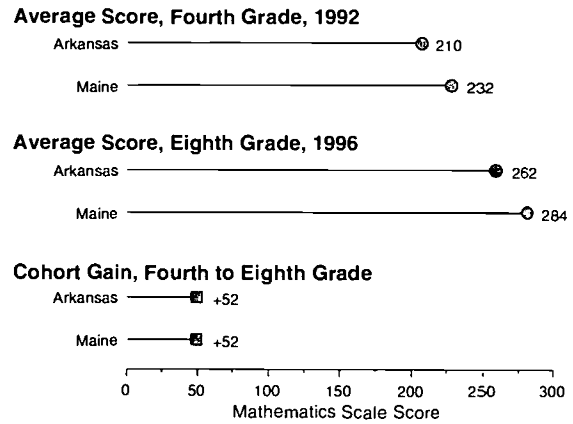
Table 5: NAEP Math Scores and Cohort Growth, by Region of the Country			
	4TH GRADE, 1992	8TH GRADE, 1996	COHORT GROWTH
Nation	220	272	+52
Northeast	224	277	+53
Southeast	211	266	+55
Central	224	277	+53
West	219	269	+50

Source for Tables 3, 4, and 5: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep

* There were statistically significant differences between students whose parents had graduated from high school and students whose parents had graduated from college.

State Comparisons

Figure 4: Average NAEP Mathematics Scores and Cohort Growth, Arkansas and Maine



Source: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep.

The biggest differences between looking at average score trends for a particular grade and looking at cohort changes can be seen in comparisons of NAEP scores among the states. The range in average performance at a grade level among the states is very large, and these average score differences are typically described as measures of the differences among the states in the quality of their education systems. For example, the average mathematics scale score for fourth-graders in top-scoring Maine in 1992 was 232, compared to 210 in bottom-scoring Arkansas. At the eighth grade, in 1996, the average was 284 in Maine, compared to 262 in Arkansas.[3]

In Figure 4 we can see how differences in these average grade level scores compare with differences in cohort growth. The cohort growth from the fourth grade (1992) to the eighth grade (1996) was 52 points in both Maine and Arkansas. While the students in the two states started at quite different levels in 1992, their cohort growth was the same, leaving them the same distance apart at grade 8 by 1996.

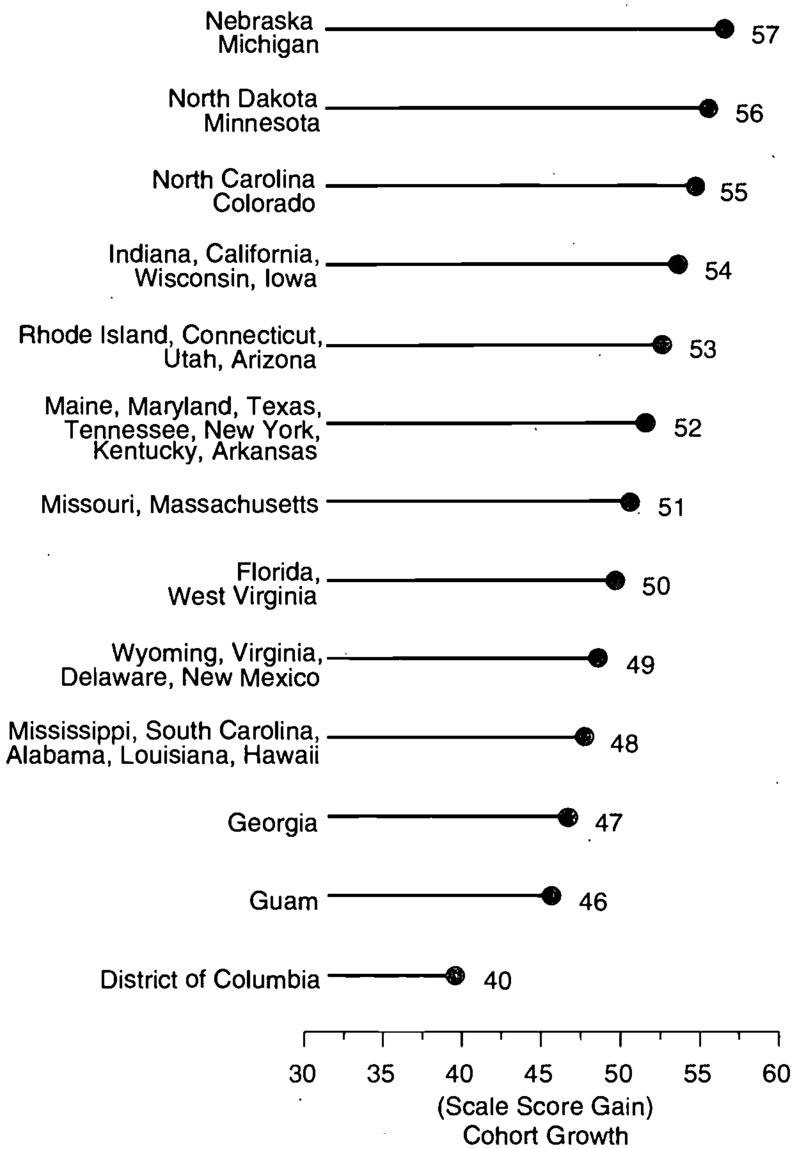
The question this type of analysis raises is whether the education system is any better in Maine than it is in Arkansas. Or, more specifically, whether Maine is doing a better job in the fifth, sixth, seventh, and eighth grades. Should we hold school systems accountable for the *level* of their students' achievement, or for the *growth* in achievement that they are able to bring about?

A few states indeed do better than others at increasing achievement in the cohort from the fourth to the eighth grade. This can be seen in Figure 5. Nebraska and Michigan do considerably better with their +57 scale points of cohort growth than the District of Columbia with +40. But 21 states out of 37 participating in NAEP in both 1992 and 1996 are clustered from +55 to +50 of cohort growth. This five-point spread is equivalent to what is learned in about four months of school.

A more precise way to compare the states is

[3] Some researchers have suggested that state NAEP scores should be adjusted to standardize for the demographic characteristics of the state. Such standardization will narrow the differences among the states. For a discussion and application of this topic see Howard Wainer and Edward Kulik, *A Comparative Study of the Academic Performance of Pennsylvania's Public School Children: Mathematics and Reading between 1990 and 1996*, Research Report 97-25, Princeton, NJ: Educational Testing Service, December 1977.

Figure 5: NAEP Mathematics Cohort Growth (from the Fourth Grade - 1992 to the Eighth Grade - 1996), by Participating States



Source: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep.

to take into account the "standard errors" which result from the fact that the data are drawn based on samples of students, rather than all the students in a state. The important question then becomes what differences among the states are statistically significant, unlikely to occur by chance. In the chart in this report's appendix, each state's mathematics cohort growth can be compared with every other state's cohort growth for significant differences. Reading it like a "mileage chart," and following Nebraska's line across the top, a "+" shows the 13 states that Nebraska exceeds by a significant difference, including Guam and the District of Columbia. Michigan does significantly better than six states, North Dakota does better than 11, Minnesota does better than eight, and Colorado outperforms six states. Dropping down the chart to West Virginia, "-" signs are shown. These signs appear for states that exceed West Virginia's performance by a significant difference.

But the major conclusion that can be drawn from the chart comes from all the empty spaces in the middle. All these empty spaces mean that there were no significant differences between the two intersecting states. *Most of the states are not significantly different from each other in terms of cohort growth from the fourth to the eighth grade.*

Describing and Understanding Cohort Growth

We have compared cohort growth over distant periods of time and among different groups of students for four-year periods in which data are available. But it is hard to visualize just what such levels of cohort growth mean in terms of academic achievement. While there is no way to make this perfectly clear, it is possible to shed some light on the matter. What we can do, in mathematics for example, is look at the kinds of problems fourth-graders are likely to be able to solve successfully, and then look at the kinds of problems the cohort of students can handle four years later. We have done this in Figure 6.

On the left side of Figure 6, examples of problems are shown at their level of difficulty on the NAEP achievement scale (from 0 to 500) for fourth-grade students in 1992. At the bottom end of the range (192), an example is "subtracting whole numbers with regrouping." This is approximately where the average Black fourth-grader scored in 1992 (193). This contrasts with the average White fourth-grader, who could do things like "represent a system algebraically," at the 227 score level. Hispanic and American Indian students are arrayed in between.

Shifting to the eighth-grade results on the right side of the chart, the average Black student scored just below the level where students can do things such as "round decimals to nearest whole numbers." White students are, on average, at a level where they can "use a pattern to draw a path on a grid," at the 282 score level. The figure shows the relative positions of fourth- and eighth-graders on the NAEP achievement scale.

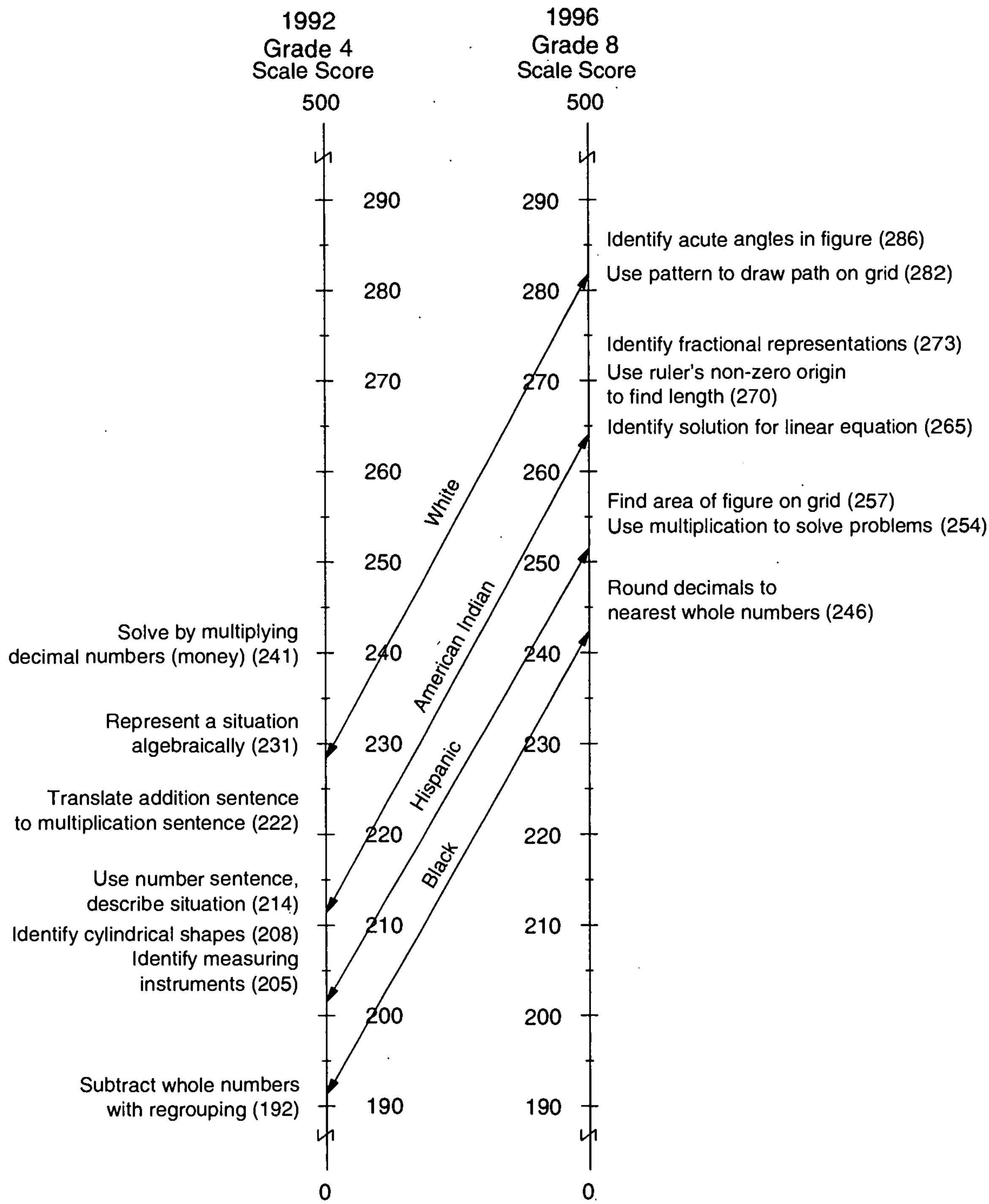
Figure 6 also shows the *growth* in achievement for fourth-graders over their next four years of school — about 50 points on the NAEP scale. And while there is considerable disparity by race/ethnicity at the fourth and eighth grades, as pointed out above, there is no statistically significant difference in terms of the amount of improvement from fourth grade to eighth grade by race/ethnicity.

For Black students, the score improvement of 50 points brought them to a point in the eighth grade where they were only slightly above the average for fourth-grade White students. The *gain* is similar, but the *level* is very different, and the examples give some palpability to what the numbers mean in terms of achievement comparisons.

Figure 7 may also be useful in explaining the progression of learning from the fourth to the eighth grade, and the levels achieved. It shows, on the left, a range of test items used in the NAEP report to illustrate difficulty levels on the NAEP scale for the 1996 mathematics assessment. On the right, the figure shows the average scale scores for both grade levels and for racial/ethnic subgroups. It also shows the proficiency levels. or standards, set by the National Assessment Governing Board.

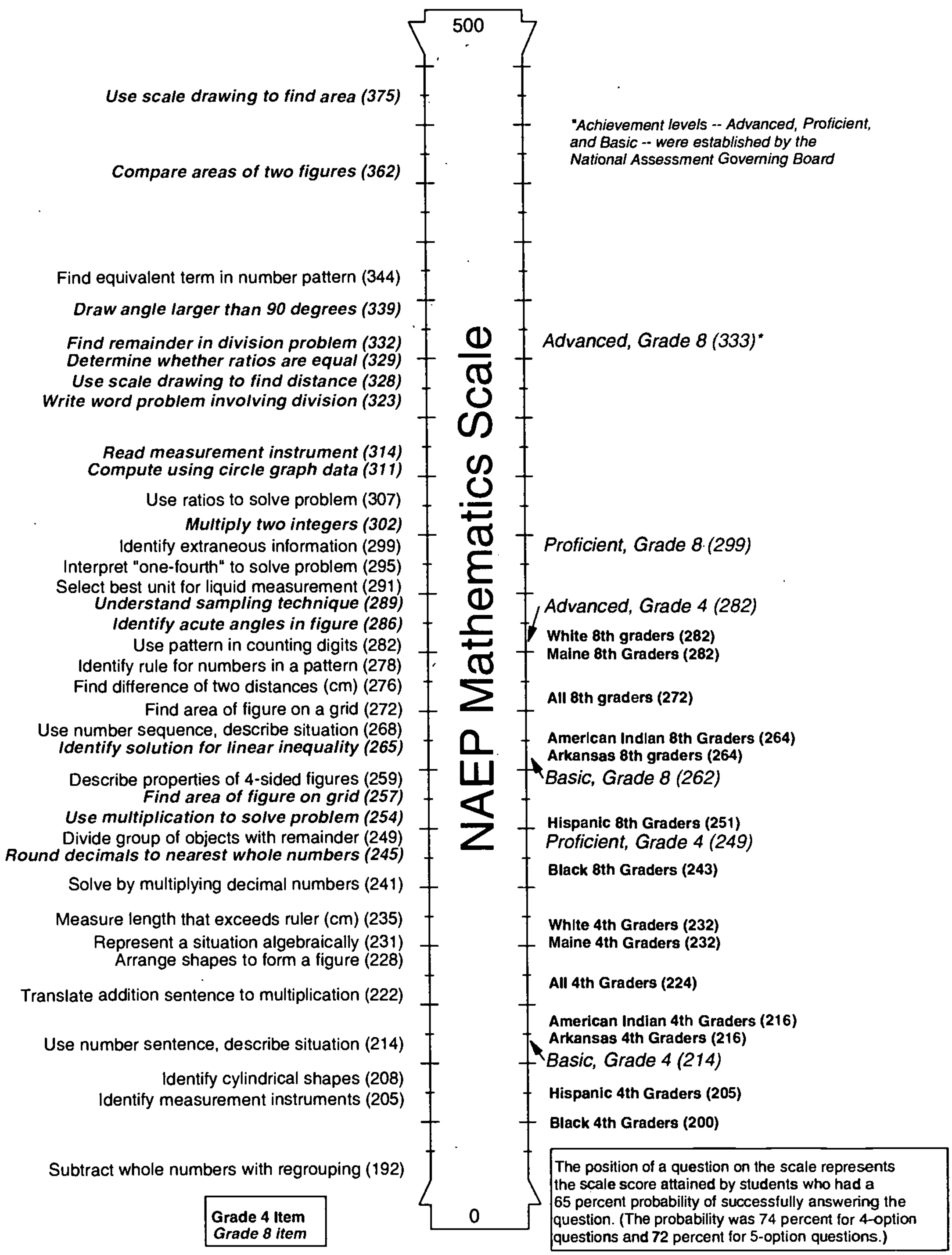
Figure 7 also shows where Maine and Arkansas, states at opposite ends of the score continuum, fall. While the questions asked of eighth-graders were not asked of fourth-graders, the array can illustrate the level of difficulty at different parts of the NAEP scale. What is striking in addition to the differences among student subgroups is how close eighth-grade Black students are to fourth-grade White students and how the "advanced level" for the fourth grade is considerably higher than the "basic level" for the eighth grade.

Figure 6: Cohort Growth in Mathematics from the Fourth to the Eighth Grade, 1992 to 1996



Source: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep.

Figure 7: Map of Selected Items on the NAEP Mathematics Scale and Average Scale Scores of Fourth- and Eighth-Graders, 1996



Source: National Assessment of Educational Progress data analyzed by the ETS Policy Information Center. See http://nces.ed.gov/naep.

In Conclusion

While in most cases the average NAEP scores of today's students are slightly higher than those of students 20 or 25 years ago, the cohort growth between the fourth and eighth grade is not. In fact, cohort growth is the same as, or lower than, it was during the earliest period for which we have data.

And when we compare states, there is little difference in the cohort growth between the fourth and eighth grade. While Maine was the top-scoring state in the nation and Arkansas was the bottom-scoring state, both states had the same cohort growth, 52 points on the NAEP scale between the fourth and eighth grade.

How do we, and how should we, look at NAEP scores in reaching a judgment as to whether the education system is performing better or worse over time? Are Maine and Arkansas at the two ends of the school quality continuum, or are they actually equal?

Average NAEP scores for a particular age or grade over time tell us whether students know more or less than their counterparts in an earlier period. This is valuable information, quite apart from the question of whether it is the best measure of *school* effectiveness. We recognize that learning occurs at home and in the community as well as in the school, and that the richness of early home and life experiences affects how well students do in school.

However, we can get closer to what actually happens in school if we focus on achievement growth while students are in school. There appears to have been no change in the cohort growth between the fourth and the eighth grade over the last 20 or 25 years except in math, where there was a decline. It is also striking, over a four-year period, how little difference there is in what schools across the states add to achievement.

It is not our intention to determine which is the best measure, for examining average score trends and cohort growth tell us different things. But it *does* appear to be important to look at *both* measures. NAEP was redesigned in 1984 to provide both views of educational achievement, by spacing grade level and assessment years four years apart, and by having a single achievement scale on which students in the three different grades (or age levels) could be placed.

However, this single developmental scale is beginning to be abandoned in favor of separate scales for each grade level. Thus, in the *NAEP 1996 Science Report Card for the Nation and the States,* each of the three grades is measured on a separate scale. It is not possible to see how students compare in knowledge among the fourth, eighth and twelfth grades, so it is not possible to look at gain in science learning in 1996. It adds an important dimension to be able to say how much students *learned,* as well as how much they *know.*

The availability of both kinds of achievement measures leads to another important question. Should performance standards be set only for achievement at a single grade level? This is the method used now by the National Assessment Governing Board, the group that sets policy for NAEP. NAEP reports the percentage of students at the Basic, Proficient, and Advanced levels at each of the three grade levels. Should standards also be set for how much achievement *growth* should be realized between the fourth and eighth grade? The top state has a gain of 57 points in mathematics between grades 4 and 8. Is that enough? The bottom state has a gain of 40 points, which is clearly not enough. But how much is enough? Maine and Arkansas each show a gain of 52 points, but one state has the highest average score in the nation and the other has the lowest. Are the student achievement gains in Maine and Arkansas where we think they should be? Basically, when sampling error is taken into account, only a few states have cohort growth that is significantly different from the rest of the states.

Another way of framing the question is to ask how high on the scale do we have to be at the eighth grade to maintain the international ranking we achieved at the fourth grade in TIMSS? What is the target we need to hit to become world-class by the eighth grade? Or what cohort score gain could be expected if the National Council of Teachers of Mathematics standards were implemented in all schools?

NAEP has been particularly important in permitting the tracking of "value added" because such measurement is almost never done elsewhere in educational testing. One exception is the Tennessee Value-Added Assessment System, in place since 1992. This system has enabled Tennessee to associate achievement with possible causes that could not be measured with traditional testing systems. One of the state's principal findings is that the largest single factor affecting academic growth is differences in the effectiveness of individual teachers.[4]

The TIMSS study has focused attention on why U.S. students slip between the fourth and the eighth grade, relative to other countries. To examine this issue closely, we need to look at the extent and pattern of growth between these grades. While NAEP reports have not addressed "value added" specifically, the data have been collected and can provide information not available from simply comparing average achievement in grade 4 or 8 with that of some prior time. This aspect of NAEP holds considerable promise. We have tried to show, in this brief report, the kinds of insights that such analysis might permit and the kinds of questions it brings to our attention.

Measuring and examining cohort growth provide a different and important dimension in understanding trends in educational achievement. Of course, such efforts provide no information about *why* students achieve or grow at different levels. Research must determine the factors related to these cohort score changes. The cohort growth changes examined in this report represent a different set of outcome variables with which researchers can work.

[4] See "The Value-Added Side of Standards," by Chris Pipho, *Phi Delta Kappan*, Vol. 79, No. 5, January 1998, p. 341.

Appendix: Growth in Mathematics Score from 1992 (Grade 4) to 1996 (Grade 8)

	Mean	SE	NEBRASKA	MICHIGAN	N DAKOTA	MINN	N. CAR	COLORADO	INDIANA	CALIFOR	WISCON	IOWA	RHODE IS	CONN	UTAH	ARIZONA	MAINE	MARYLAND	TEXAS	TENN	NEW YORK	KENTUCKY	ARKANSAS	MISSOURI	MASS	FLORIDA	WEST VA	WYOMING	VIRGINIA	DELAWARE	NEW MEX	MISSISS	S. CAR	ALABAMA	LOUISIAN	HAWAII	GEORGIA	GUAM	DC
NEBRASKA	57.44	1.60																									+	+	+	+	+	+	+	+	+	+	+	+	+
MICHIGAN	56.99	2.47																												+		+				+	+	+	+
N DAKOTA	55.56	1.19																									+	+		+	+	+	+		+	+	+	+	+
MINN	55.56	1.62																										+		+		+	+			+	+	+	+
N. CAR	54.95	1.80																																			+	+	+
COLORADO	54.59	1.46																												+		+				+	+	+	+
INDIANA	54.49	1.78																																				+	+
CALIFOR	54.37	2.42																																					+
WISCON	54.16	1.87																																				+	+
IOWA	54.10	1.66																																				+	+
RHODE IS	53.43	1.79																																				+	+
CONN	52.79	1.59																																				+	+
UTAH	52.73	1.41																																				+	+
ARIZONA	52.62	1.89																																					+
MAINE	52.42	1.63																																					+
MARYLAND	52.36	2.49																																					+
TEXAS	52.28	1.88																																					+
TENN	52.17	1.94																																					+
NEW YORK	51.79	2.08																																					+
KENTUCKY	51.54	1.47																																					+
ARKANSAS	51.45	1.76																																					+
MISSOURI	51.07	1.82																																					+
MASS	50.97	2.10																																					+
FLORIDA	49.95	2.37																																					+
WEST VA	49.60	1.47	-		-																																		+
WYOMING	49.40	1.30	-		-	-																																	+
VIRGINIA	48.99	2.03	-																																				+
DELAWARE	48.83	1.23	-	-	-	-		-																															+
NEW MEX	48.67	1.89	-		-																																		+
MISSISS	48.39	1.61	-	-	-	-		-																															+
S. CAR	48.28	1.88	-		-	-																																	+
ALABAMA	48.27	2.66	-																																				
LOUISIAN	48.24	2.15	-		-																																		+
HAWAII	48.07	1.63	-	-	-	-		-																															+
GEORGIA	46.87	2.06	-	-	-	-	-	-																															
GUAM	45.80	1.87	-	-	-	-	-	-	-		-	-	-	-	-																								
DC	40.23	1.45	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-	-			

21

04202-14801 • DY58M7 • 204862 • Printed in U.S.A.

NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").